

## CLAIMS

1           1. (currently amended) A method for synthesizing an auditory scene, comprising the steps  
2       of:  
3           (a) dividing an input audio signal into a plurality of different frequency bands; and  
4           (b) applying two or more different sets of one or more spatial parameters to two or more of  
5       the different frequency bands in the input audio signal to generate two or more synthesized audio signals  
6       of the auditory scene, ~~wherein for each of the two or more different frequency bands, the corresponding~~  
7       ~~set of one or more spatial parameters is applied to the input audio signal as if the input audio signal~~  
8       ~~corresponded to a single audio source in the auditory scene.~~

1           2. (currently amended) The invention of claim 1, wherein:  
2       the input audio signal corresponds to a combination of audio signals from two or more different  
3       audio sources; and  
4        each set of one or more spatial parameters corresponds to a different audio source in the auditory  
5       scene.

1           3. (original) The invention of claim 1, wherein, for at least one of the sets of one or more  
2       spatial parameters, at least one of the spatial parameters corresponds to a combination of two or more  
3       different audio sources in the auditory scene that takes into account relative dominance of the two or  
4       more different audio sources in the auditory scene.

1           4. (original) The invention of claim 1, wherein the input audio signal is a mono signal.

1           5. (currently amended) The invention of claim [[4]] 1, wherein the ~~mono~~ input audio signal  
2       corresponds to a combination of two or more different ~~mono~~ source signals, wherein the two or more  
3       different frequency bands are selected by comparing magnitudes of the two or more different ~~mono~~  
4       source signals, wherein, for each of the two or more different frequency bands, one of the ~~mono~~ source  
5       signals dominates the other ~~mono~~ source signals.

1           6. (currently amended) The invention of claim [[4]] 1, wherein the ~~mono~~ input audio signal  
2       corresponds to a combination of left and right audio signals ~~of a binaural signal~~, wherein each different  
3       set of one or more spatial parameters is generated by comparing the left and right audio signals in a  
4       corresponding frequency band.

1           7. (original) The invention of claim 1, wherein step (a) comprises the step of dividing the  
2       input audio signal into the plurality of different frequency bands based on information corresponding to  
3       the different sets of one or more spatial parameters.

1           8. (currently amended) The invention of claim 1, wherein:  
2       the input audio signal corresponds to a combination of audio signals from two or more different  
3       audio sources; and  
4        each set of one or more spatial parameters is applied to at least one frequency band in which the  
5       input audio signal is dominated by a corresponding audio source in the auditory scene.

1           9. (original) The invention of claim 1, wherein each set of one or more spatial parameters  
2       comprises one or more of an interaural level difference, an interaural time delay, and a head-related  
3       transfer function.

1           10. (original) The invention of claim 1, wherein:  
2           step (a) further comprises the step of converting the input audio signal from a time domain into a  
3           frequency domain; and  
4           step (b) further comprises the step of converting the two or more synthesized audio signals from  
5           the frequency domain into the time domain.

1           11. (currently amended) The invention of claim 1, wherein the two or more synthesized  
2           audio signals comprise left and right audio signals ~~of a binaural signal~~ corresponding to the auditory  
3           scene.

1           12. (currently amended) The invention of claim 1, wherein the two or more synthesized  
2           audio signals comprise [[two]] three or more signals of a multi-channel audio signal corresponding to the  
3           auditory scene.

1           13. (currently amended) The invention of claim 1, wherein:  
2           the input audio signal is a mono signal;  
3           each set of one or more spatial parameters corresponds to a different audio source in the auditory  
4           scene;  
5           step (a) comprises the steps of:  
6               (1) converting the mono signal from a time domain into a frequency domain;  
7               (2) dividing the converted mono signal into the plurality of different frequency  
8           bands based on information corresponding to the sets of one or more spatial parameters;  
9               each set of one or more spatial parameters is applied to at least one frequency band in which the  
10          input audio signal is dominated by a corresponding audio source in the auditory scene;  
11          each set of one or more spatial parameters comprises one or more of an interaural level  
12          difference, an interaural time delay, and a head-related transfer function;  
13          the two or more synthesized audio signals comprise left and right audio signals ~~of a binaural~~  
14          signal corresponding to the auditory scene; and  
15           step (b) further comprises the step of converting the left and right audio signals from the  
16          frequency domain into the time domain.

1           14. (original) The invention of claim 13, wherein the mono signal corresponds to a  
2           combination of two or more different mono source signals, wherein the two or more different frequency  
3           bands are selected by comparing magnitudes of the two or more different mono source signals, wherein,  
4           for each of the two or more different frequency bands, one of the mono source signals dominates the  
5           other mono source signals.

1           15. (currently amended) The invention of claim 13, wherein the mono signal corresponds to  
2           a combination of left and right audio signals ~~of a binaural signal~~, wherein each different set of one or  
3           more spatial parameters is generated by comparing the left and right audio signals in a corresponding  
4           frequency band.

1           16. (currently amended) A machine-readable medium, having encoded thereon program  
2           code, wherein, when the program code is executed by a machine, the machine implements a method for  
3           synthesizing an auditory scene, comprising the steps of:  
4               (a) dividing an input audio signal into a plurality of different frequency bands; and  
5               (b) applying two or more different sets of one or more spatial parameters to two or more of  
6           the different frequency bands in the input audio signal to generate two or more synthesized audio signals  
7           of the auditory scene, ~~wherein for each of the two or more different frequency bands, the corresponding~~

8 set of one or more spatial parameters is applied to the input audio signal as if the input audio signal  
9 corresponded to a single audio source in the auditory scene.

1 17. (currently amended) An apparatus for synthesizing an auditory scene, comprising:

2 (a) means for dividing an input audio signal into a plurality of different frequency bands;  
3 and

4 (b) means for applying two or more different sets of one or more spatial parameters to two or  
5 more of the different frequency bands in the input audio signal to generate two or more synthesized audio  
6 signals of the auditory scene, wherein for each of the two or more different frequency bands, the  
7 corresponding set of one or more spatial parameters is applied to the input audio signal as if the input  
8 audio signal corresponded to a single audio source in the auditory scene.

1 18. (currently amended) An apparatus for synthesizing an auditory scene, comprising:

2 (1) an auditory scene synthesizer configured to:

3 (a) divide an input audio signal into a plurality of different frequency bands; and  
4 (b) apply two or more different sets of one or more spatial parameters to two or

5 more of the different frequency bands in the input audio signal to generate two or more synthesized audio  
6 signals of the auditory scene, wherein for each of the two or more different frequency bands, the  
7 corresponding set of one or more spatial parameters is applied to the input audio signal as if the input  
8 audio signal corresponded to a single audio source in the auditory scene; and

9 (2) one or more inverse time-frequency transformers configured to convert the two or more  
10 synthesized audio signals from a frequency domain into a time domain.

1 19. (currently amended) A method for processing two or more input audio signals,  
2 comprising the steps of:

3 (a) converting the two or more input audio signals from a time domain into a frequency  
4 domain;

5 (b) generating a set of one or more auditory scene parameters for each of two or more  
6 different frequency bands in the two or more converted input audio signals, where each set of one or  
7 more auditory scene parameters is generated as if the corresponding frequency band corresponded to a  
8 single audio source in an auditory scene; and

9 (c) combining the two or more input audio signals to generate a combined audio signal.

1 20. (original) The invention of claim 19, wherein:

2 the two or more input audio signals are mono signals corresponding to different audio sources in  
3 the auditory scene;

4 each set of one or more auditory scene parameters corresponds to an audio source that dominates  
5 the other audio sources in the corresponding frequency band; and

6 the two or more input audio signals are combined in the time domain to generate the combined  
7 audio signal.

1 21. (currently amended) The invention of claim 19, wherein:

2 the two or more input audio signals are left and right audio signals of a binaural signal;

3 each set of one or more auditory scene parameters is generated by comparing the left and right  
4 audio signals in the corresponding frequency band;

5 the combined audio signal is generated by performing auditory scene removal on the left and  
6 right audio signals in the frequency domain based on the two or more sets of one or more auditory scene  
7 parameters; and

8 further comprising the step of converting the combined audio signal from the frequency domain  
9 into the time domain.

1           22. (currently amended) A machine-readable medium, having encoded thereon program  
2 code, wherein, when the program code is executed by a machine, the machine implements a method for  
3 processing two or more input audio signals, comprising the steps of:

- 4           (a) converting the two or more input audio signals from a time domain into a frequency  
5 domain;  
6           (b) generating a set of one or more auditory scene parameters for each of two or more  
7 different frequency bands in the two or more converted input audio signals, ~~where each set of one or~~  
8 ~~more auditory scene parameters is generated as if the corresponding frequency band corresponded to a~~  
9 ~~single audio source in an auditory scene;~~ and  
10          (c) combining the two or more input audio signals to generate a combined audio signal.

1           23. (currently amended) An apparatus for processing two or more input audio signals,  
2 comprising:

- 3           (a) means for converting the two or more input audio signals from a time domain into a  
4 frequency domain;  
5           (b) means for generating a set of one or more auditory scene parameters for each of two or  
6 more different frequency bands in the two or more converted input audio signals, ~~where each set of one or~~  
7 ~~more auditory scene parameters is generated as if the corresponding frequency band corresponded to a~~  
8 ~~single audio source in an auditory scene;~~ and  
9           (c) means for combining the two or more input audio signals to generate a combined audio  
10 signal.

1           24. (currently amended) An apparatus for processing two or more input audio signals,  
2 comprising:

- 3           (a) a time-frequency transformer configured to convert the two or more input audio signals  
4 from a time domain into a frequency domain;  
5           (b) an auditory scene parameter generator configured to generate a set of one or more  
6 auditory scene parameters for each of two or more different frequency bands in the two or more  
7 converted input audio signals, ~~where each set of one or more auditory scene parameters is generated as if~~  
8 ~~the corresponding frequency band corresponded to a single audio source;~~ and  
9           (c) a combiner configured to combine the two or more input audio signals to generate a  
10 combined audio signal.

1           25. (original) The invention of claim 24, wherein:

- 2           the two or more input audio signals are mono signals corresponding to different audio sources in  
3 the auditory scene;  
4           each set of one or more auditory scene parameters corresponds to an audio source that dominates  
5 the other audio sources in the corresponding frequency band; and  
6           the combiner operates in the time domain.

1           26. (currently amended) The invention of claim 24, wherein:

- 2           the two or more input audio signals are left and right audio signals ~~of a binaural signal;~~  
3           each set of one or more auditory scene parameters is generated by comparing the left and right  
4 audio signals in the corresponding frequency band;  
5           ~~the combiner is configured to perform auditory scene removal on the left and right audio signals~~  
6 ~~in the frequency domain based on the two or more sets of one or more auditory scene parameters;~~ and  
7           further comprising an inverse time-frequency transformer configured to convert the combined  
8 audio signal from the frequency domain into the time domain.

1           27. (new) The invention of claim 1, wherein, for each of the two or more different  
2 frequency bands, the corresponding set of one or more spatial parameters is applied to the input audio  
3 signal as if the input audio signal corresponded to a single audio source in the auditory scene.

1           28. (new) The invention of claim 1, wherein the input audio signal corresponds to a  
2 combination of three or more audio signals of a multi-channel signal, wherein each different set of one or  
3 more spatial parameters is generated by comparing at least two of the audio signals in a corresponding  
4 frequency band.

1           29. (new) The invention of claim 1, further comprising decompressing a compressed audio  
2 signal to generate the input audio signal.

1           30. (new) The invention of claim 19, wherein each set of one or more auditory scene  
2 parameters is generated as if the corresponding frequency band corresponded to a single audio source in  
3 an auditory scene.

1           31. (new) The invention of claim 19, wherein:  
2           the two or more input audio signals are three or more audio signals of a multi-channel signal; and  
3           each set of one or more auditory scene parameters is generated by comparing at least two of the  
4 audio signals in the corresponding frequency band.

1           32. (new) The invention of claim 19, further comprising compressing the combined audio  
2 signal to generate a compressed audio signal.

1           33. (new) The invention of claim 19, wherein the combined audio signal is generated by  
2 performing auditory scene removal on the input audio signals in the frequency domain based on the two  
3 or more sets of one or more auditory scene parameters.

1           34. (new) The invention of claim 19, wherein the combined audio signal is generated by  
2 averaging the input audio signals.

1           35. (new) A bitstream comprising a combined audio signal and a plurality of auditory scene  
2 parameters, wherein:  
3           the combined audio signal is generated by combining two or more input audio signals; and  
4           the auditory scene parameters are generated by:  
5           converting the two or more input audio signals from a time domain into a frequency  
6 domain; and  
7           generating a set of one or more auditory scene parameters for each of two or more  
8 different frequency bands in the two or more converted input audio signals.